Using a radial ultrasound probe's virtual origin to compute midsagittal smoothing splines in polar coordinates

Matthias Heyne[1,a)] **and Donald Derrick**[2]
[1]*Department of Linguistics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand*
[2]*New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Private Bag 4800, Christchurch, New Zealand*
matthias.heyne@pg.canterbury.ac.nz, donald.derrick@canterbury.ac.nz

Abstract: Tongue surface measurements from midsagittal ultrasound scans are effectively arcs with deviations representing tongue shape, but smoothing-spline analysis of variances (SSANOVAs) assume variance around a horizontal line. Therefore, calculating SSANOVA average curves of tongue traces in Cartesian Coordinates [Davidson, J. Acoust. Soc. Am. **120**(1), 407–415 (2006)] creates errors that are compounded at tongue tip and root where average tongue shape deviates most from a horizontal line. This paper introduces a method for transforming data into polar coordinates similar to the technique by Mielke [J. Acoust. Soc. Am. **137**(5), 2858–2869 (2015)], but using the virtual origin of a radial ultrasound transducer as the polar origin—allowing data conversion in a manner that is robust against between-subject and between-session variability.

[AL]
Date Received: July 16, 2015 **Date Accepted:** November 23, 2015

1. Introduction

Calculating smoothing-spline analysis of variance (SSANOVA) average curves (Gu, 2002; Wang, 2011; Davidson, 2006) of midsagittal ultrasound tongue contours based on Cartesian coordinates has recently been shown to create errors for any sections of the tongue that are not parallel to the horizontal axis (Mielke, 2015). However, Mielke's technique estimates the origin of the polar coordinate system based on extreme values from the collected tongue surface traces (see Mielke, 2015, p. 2861). This paper introduces a method of estimating the polar origin from a radial transducer's theoretical point of origin for ultrasound waves, or virtual origin. Because our method of polar coordinate conversion is based on the design of radial ultrasound probes, it allows grounded statistical comparison of SSANOVA average curves for data recorded using different radial probes or originating from different participants. It also allows for identification of the probe's origin independent of the quality of the imaged tongue surface, or how much of the tongue tip and root were captured. Furthermore, this method allows rotations to align the tongues of different participants that mimic the tongue shape alignment benefit of Mielke's method for identifying the polar origin, but without risking the introduction of subtle errors based on the difference between the physical origin of the ultrasound waves and the assumed origin of the polar coordinates.

2. Reasons for errors resulting from calculating SSANOVA curves in the Cartesian plane

The comparison of SSANOVA average curves based on individual tongue traces expressed in Cartesian coordinates operates on the dimensions tongue height (y-axis) and anteriority (x-axis; cf. Mielke, 2015), and introduces two errors into tongue measurements: The first stems from the fact that the "tongue surface typically approximates an arc more closely than it approximates a horizontal line" (Mielke, 2015, p. 2861). Variation around the arc represents tongue shape and measurement error, but SSANOVA assumes variation around a horizontal line. As measurement extends toward the tongue tip or root, the actual variance is around an increasingly vertical slope; the mismatch generates systematic analysis errors. The second error is that the Cartesian coordinates are not compatible with the physical reality of ultrasound data collection when using a radial transducer. Such a transducer sends out ultrasonic rays

a)Author to whom correspondence should be addressed. Also at: New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Private Bag 4800, Christchurch, New Zealand.

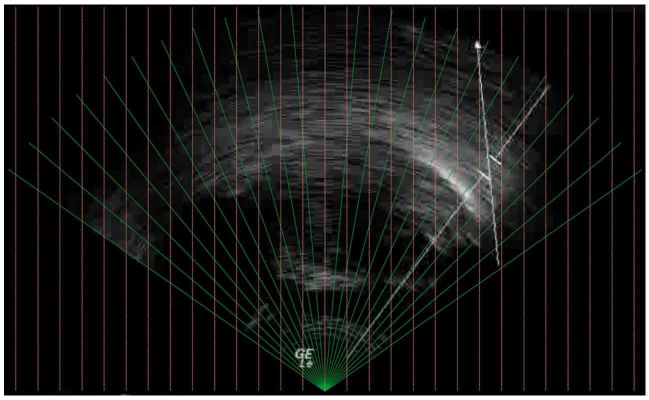

Fig. 1. (Color online) Lines superimposed on an ultrasound image, illustrating the incorrect assumption underlying the use of Cartesian coordinates when calculating SSANOVA average curves (vertical lines), and the sweeping nature of ultrasound data collection (fan-shaped lines).

sequentially in radial fashion covering a wedge-shaped scan area, which means that all measurements are taken along radial lines extending from a virtual origin point for the probe. The result of both is that the position of the mean tongue surface is incorrect, and the variance around that mean is larger than it should be, worsening at the tongue tip and root. Thus, when calculating an SSANOVA average curve based on Cartesian coordinates, the resulting average curve and error bounds do not correctly represent the variation in the data at that particular location.

Figure 1 illustrates the sweeping nature of data collection (fan-shaped lines) as compared to the incorrect assumption underlying the use of Cartesian coordinates (vertical lines). It can be easily seen that the discrepancy between the mode of data collection and the assumption underlying the calculation of an SSANOVA average curve in the Cartesian plane is most pronounced at the edges of the scan area, which is where the most significant errors have been shown to occur (Mielke, 2015).

When the coordinates of individual tongue traces are instead transformed to the polar plane and thus rendered relative to the physical position of the transducer, the variation in the data is represented much more accurately, both because of the nature of the method of data collection and the general elliptical shape of a tongue at rest. Figure 2 allows the comparison of an SSANOVA average curve of 260 repetitions of the vowel /ə/ produced by a speaker of New Zealand English calculated in the Cartesian plane (dashed line) with a corrected representation based on polar coordinates relative to the virtual origin (solid line); notice that the transformation to polar coordinates also reduces the error estimates, printed as dotted lines.

3. An alternative way of choosing the polar coordinate system origin

We agree with Mielke's motivation for using polar coordinates based on tongue shape, and doing so eliminates the first error. But we argue the solution to the second error is to take into account the nature of ultrasound data collection when choosing an origin for the transposition of Cartesian coordinates to the polar plane. In ultrasound image recordings, the quality of the image may be poor, causing only part of the tongue

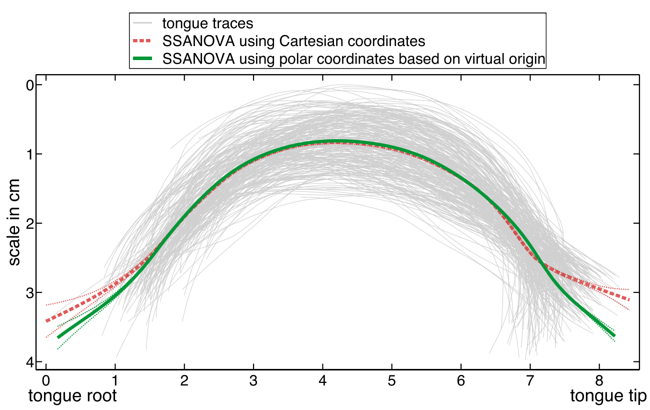

Fig. 2. (Color online) Tongue traces of 260 repetitions of the vowel /ə/ produced by a speaker of New Zealand English (faint lines), Cartesian SSANOVA (dashed line), polar SSANOVA (solid line). The dotted lines represent 95% confidence intervals.

surface to be visible. In addition, part of the tongue tip or root may be systematically occluded depending on the speech articulation in question.

Automatically selecting a suitable polar coordinate origin based solely on extreme values in the collected traces as in the paper by Mielke (2015), means giving up the only fixed point for data collected with an imaging method that does not provide any anatomical landmarks (cf. Stone, 2005).

3.1 Considerations

Ultrasound images are produced by tracking ultrasound emitted from piezo-electric crystals inset into a transducer and reflected back when they encounter a change in density, such as the air-boundary above the surface of the tongue. No single piezo-electric crystal sends out and receives the echo from all of the ultra-high-frequency sound waves that go into creating a single ultrasound frame—it is actually a rather large number of individual crystals.

With linear transducers, there is no single virtual origin point of the ultrasound waves. Therefore, approximating polar coordinates based on the tongue shape can still address the first source of error—that tongue shapes themselves are variants of an arc, whereas SSANOVA tests assume variation based on a horizontal line. For these transducers, Mielke's technique is better than using Cartesian coordinates, providing a system for estimating the rectification of tongue shape data with regard to these SSANOVA assumptions.

However, in radial ultrasound probes, the crystals are aligned such that the ultrasound waves can be treated as though they emanate from a single point. Selecting this point for the origin of the polar coordinates will best represent the mode of ultrasound data collection from a radial transducer. Our modification of Mielke's technique is thus intended for use only with radial ultrasound transducers, and it exploits their main advantage, which is providing a virtual origin that remains constant for similar probes and across subjects.

3.2 Estimating the virtual origin from an ultrasound frame

In order to estimate the ultrasound transducer origin from a randomly selected individual ultrasound frame, we plotted two lines (outside fan lines in Fig. 1) along the outside of the scan area and set their equations equal to approximate the implied transducer origin which is not directly visible on the ultrasound image.

Polar coordinates express points in two-dimensional space as an angle θ relative to a reference direction, and a distance r to a particular reference point. For our application, it was easiest to use a vertical line emanating from the virtual origin as the reference direction and to calculate Euclidian distances from the virtual origin defined as $O(x_O, y_O)$ to any point on a tongue trace defined as $P(x_P, y_P)$. The following equations [Eqs. (1) and (2)] resolve to translate the Cartesian coordinates of P to the polar plane (with the virtual origin as vertex):

$$r_P = \sqrt{(x_P - x_O)^2 + (y_P - y_O)^2}, \tag{1}$$

$$\theta_P = \tan^{-1}\left(\frac{(y_P - y_O)}{(x_P - x_O)}\right). \tag{2}$$

Our approach of plotting lines on an ultrasound image to estimate the virtual original also allows the translation of the pixel distances forming the basis of the exported contours' coordinates into the centimeter scale displayed in the ultrasound video.

3.3 Measurement source independent of measurement quality

Our approach provides a polar origin that does not change based on how much or what portion of the tongue surface was captured. Depending on the speech act in question, quality of the tongue image (which differs from speaker to speaker), or shape of a speaker's vocal tract, parts of the tongue tip or back can be absent from an ultrasound recording. For example, current research on co-collection and co-registration of ultrasound and electromagnetic articulometry (EMA) supports the longstanding notion that the tongue tip may not always be visible in ultrasound images of the tongue. Note that EMA is a data collection system using wired sensors in an electromagnetic field to track flesh points inside the body (Derrick *et al.*, 2015).

In Fig. 3, we show the average results of ten tokens of an American English speaker producing the word "auditor," overlaying ultrasound traces of the tongue and

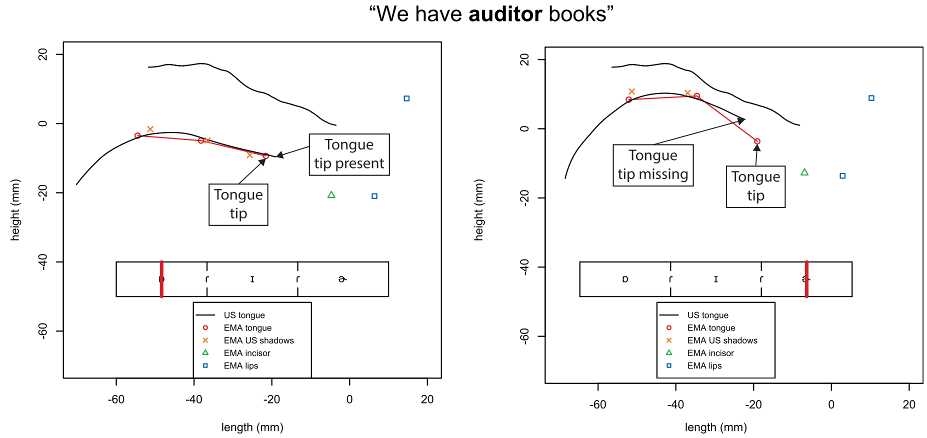

Fig. 3. (Color online) Image of ultrasound tongue contour overlayed on articulometry sensors showing that the ultrasound image of the tongue tip is present with non-rhotic vowels (example left), but occluded with rhotic vowels (example right). Palate trace acquired using EMA and ultrasound contour of the tongue surface (black lines), articulometry sensors on the tongue surface (circles), shadows of articulometry sensors in the ultrasound images (×'s), articulometry at lower incisor (triangle), articulometry at the lips (squares).

EMA data (along with the approximate location of ultrasound shadows of the EMA sensors). While the tongue tip is visible as part of the ultrasound trace at the vocalic center of a non-rhotic vowel, overlapping with an EMA sensor 0.5 cm away from the tongue tip (right-most circle), it is missing from the trace at the vocalic center of the rhotic vowel (circle to the right of and lower than end of ultrasound trace). The co-registration portion of the joint ultrasound/EMA research is currently in progress, so we provide the image for illustrative purposes only. Nevertheless, it shows that ultrasound can capture different portions of the tongue surface depending on context, and this shows a potential risk of relying on ultrasound tongue traces for estimating a polar origin.

3.4 Comparison of average contours calculated using both approaches

That the position of the polar origin chosen to transform Cartesian coordinates to the polar plane matters is illustrated in Fig. 4: When comparing our "virtual origin" and Mielke's "automatic estimation" approach, small discrepancies in curve shape and error estimates arise at the edges of the average curves. Note that we used the values of traces for all vowels in our New English speaker's vowel system to estimate the polar origin for Mielke's automatic estimation method, although only three sounds and their respective confidence intervals are shown in Fig. 4 for the sake of clarity.

The differences between the average curves in Fig. 4 show that the chosen origin affects curve shape and error estimates (dotted lines) for different types of curves in different ways—while the average curves and error estimates are fairly close for the sounds /a/ and /ʊ/, they differ more markedly for /iː/. Our method of using the virtual

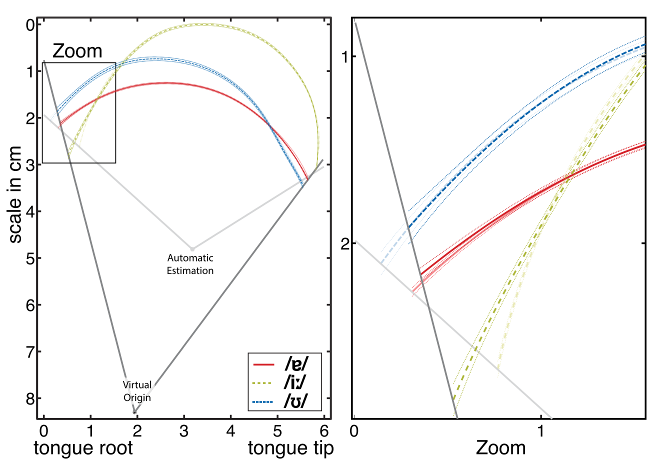

Fig. 4. (Color online) SSANOVA average curves for selected vowels of New Zealand English, produced by a single speaker during a single recording session. Faint lines (averages calculated using Mielke's approach), Dark lines (averages calculated using the virtual origin). The dotted lines represent 95% confidence intervals.

origin therefore provides a way of reducing some of the problems of the ultrasound technique pertaining to between-session and between-subject comparability by eliminating the arbitrariness associated with an automatic estimation approach.

It is also worthy of note that our approach makes precisely the same assumptions as overlaying a concentric grid on an ultrasound image as implemented in the AAA software (Articulate Instruments, 2010) and Ultra-CATS (Gu *et al.*, 2004), although it greatly enhances the statistical resolution (comparisons in AAA and Ultra-CATS are limited to a maximum of 42 and 35 fan lines, respectively; cf. Lawson *et al.*, 2013; Gu *et al.*, 2004).

4. Rotating tongue contours

By using the approximate physical position of the transducer as polar coordinate origin, it also becomes possible to rotate individual traces and/or average curves without affecting the true variation underlying their error estimates.

Since we know that "tissue edges perpendicular to the ultrasound beam are imaged best and those approaching a parallel to the beam orientation are imaged most poorly" (Stone, 2005, p. 461), rotating data relative to a vertex that does not represent the origin of the ultrasonic rays used to make these measurements would render these error estimates inappropriate at their new location.

Correct rotation employing the virtual origin can be effected by adding to or subtracting from the angular coordinate of each point. We have used this approach to adjust for gross misalignments that resulted from a participant touching the probe holder and thus changing the position of the non-metal jaw brace we used for probe stabilization (see Derrick *et al.*, 2015). In this case, we had adjusted the rotation to align the average curve shapes for each token type using the mean shapes, but we are working on implementing the technique for the automatic correction of co-collected and co-registered data using ultrasound imaging and electromagnetic articulography (EMA). In principle, such rotations can be used to align the tongue contours of different participants based upon independently motivated principles—the only errors thereby arising from the motives themselves, and not also from mismatches between the selected polar origin and the virtual origin.

5. Discussion

It is important to note that using the virtual origin is not as effective if the virtual origin has not been properly identified. As described above, this problem occurs if the ultrasound probe holder's position is compromised during a recording session. In these cases, position and rotation correction can be estimated from the recorded data itself; however, loss of information about the virtual origin can happen for automatically head-corrected data. Note, though, that both with the Palatron (Mielke *et al.*, 2005) and Corrected High-speed Anchored Ultrasound with Software Alignment (CHAUSA) techniques (which uses the head movement correction from Palatron; Miller and Finch, 2011), virtual origin information is always preserved. Other forms of automated head correction would require a consistent or recoverable virtual origin for each ultrasound frame because without this information it is not possible to estimate the virtual origin from the data and errors could be introduced via the translation and rotation used for motion correction. It is also very important to use the same type of ultrasound data collection procedure for each participant in order to take full advantage of the benefits of our technique for between-subject comparison—the ultrasound probe must be in approximately the same relative vocal tract position for each participant for maximum effectiveness.

Furthermore, thin radial ultrasound probes are preferable to larger ones, not just because they image more of the tongue surface, but also because they keep the tongue surface more perpendicular to the origin. Using a larger probe and/or an acoustical stand-off adds to the distance between the probe location and the tongue, potentially reducing the effectiveness of our technique. These observations lead to another concern—comparing different subjects who were recorded with different ultrasound probes introduces a source of error in our technique that is difficult to estimate, let alone correct.

Finally, it is worth noting that our technique is not intended to minimize variance. Using the virtual origin will often leave larger error estimations at the tongue tip and root than that result from using the automatic estimation approach (see Fig. 4). However, given the physical reality of the ultrasound technique, we expect larger errors to occur in these locations due to the tongue surface being non-perpendicular to the probe location (Stone, 2005, p. 461). Our technique therefore has a better chance of preserving the underlying error for more accurate analysis.

Given these considerations, transposing ultrasound midsagittal tongue contours to the polar plane before calculating SSANOVA average curves corrects for errors resulting from incorrect assumptions concerning the collection of ultrasound data. Taking advantage of the virtual origin of radial ultrasound transducers requires the estimation of the physical position of the ultrasound transducer, which can be done by plotting angled lines on any single ultrasound frame. Determining the polar coordinate origin in this way allows for making more accurate and precise between-subject and between-session data analysis and additionally permits the rotation of traces and/or average curves without otherwise affecting statistical mean or error estimates.

Acknowledgments

We would like to thank all of our participants for providing the data that allowed us to notice the shortcomings of calculating SSANOVA average curves in the Cartesian plane, the department of Phonetics at the University of Trier in Germany for letting us collect data on an ultrasound machine different from ours, the New Zealand Institute of Language Brain and Behaviour for supplying the funds to publish, Jennifer Hay for providing valuable commentary on earlier versions of this paper, and an anonymous reviewer for providing excellent suggestions on improving our manuscript.

References and links

Articulate Instruments Ltd. (**2010**). "Articulate Assistant user guide (version 1.18)" (Articulate Instruments Ltd., Edinburgh, UK).

Davidson, L. (**2006**). "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance," J. Acoust. Soc. Am. **120**(1), 407–415.

Derrick, D., Best, C. T., and Fiasson, R. (**2015**). "Non-metallic ultrasound probe holder for co-collection and co-registration with EMA," in *Proceedings of the 18th International Congress of Phonetic Sciences* (The University of Glasgow Glasgow, UK), Paper No. 0119.1-5.

Gu, C. (**2002**). *Smoothing Spline ANOVA Models* (Springer, New York).

Gu, J., Bressmann, T., Cannons, K., and Wong, W. (**2004**). *The Ultrasonographic Contour Analyzer for Tongue Surfaces (Ultra-CATS)* (University of Toronto, Toronto, ON, Canada).

Lawson, E., Scobbie, J. M., and Stuart-Smith, J. (**2013**). "Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation," J. Phonetics **41**, 198–210.

Mielke, J. (**2015**). "An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons," J. Acoust. Soc. Am. **137**(5), 2858–2869.

Mielke, J., Baker, A., Archangeli, D., and Racy, S. (**2005**). "Palatron: A technique for aligning ultrasound images of the tongue and palate," Coyote Papers: Working Papers in Linguistics, Linguistic Theory (University of Arizona, Tucson, AZ), Vol. 14, pp. 97–108.

Miller, A. L., and Finch, K. B. (**2011**). "Corrected high-frame rate anchored ultrasound with software alignment," J. Speech Language Hearing Res. **54**(2), 471–486.

Stone, M. (**2005**). "A guide to analysing tongue motion from ultrasound images," Clin. Linguistics Phonetics **19**(6-7), 455–501.

Wang, Y. (**2011**). *Smoothing Splines: Methods and Applications* (Chapman and Hall, Boca Raton, FL).